Process for preparing biotin

The invention relates to a gene construct which contains an S-adenosylmethionine synthase gene, having the sequence SEQ ID No. 1, and a biotin biosynthesis gene bioS1, bioS2 and/or bioS3, having the sequences SEQ ID No.3, SEQ ID No.5 and SEQ ID No.7, respectively, and, where appropriate, at least one further biotin synthesis gene sequence selected from the group bioA, bioB, bioF, bioC, bioD, bioH, bioP, bioW, bioX, bioY or bioR. The invention furthermore relates to organisms which contain this gene construct and to the use of the gene construct for preparing biotin, and also to a process for preparing biotin.

As a coenzyme, biotin (Vitamin H) plays an essential role in enzyme-catalyzed carboxylation and decarboxylation reactions. Biotin is consequently an essential factor in living cells. Almost all animals and some microorganisms have to take up biotin from the exterior since they are unable to synthesize biotin themselves. Biotin is therefore an essential vitamin for these organisms. By contrast, bacteria, yeasts and plants are able themselves to synthesize biotin from precursors (Brown et al. Biotechnol. Genet. Eng. Rev. 9, 1991: 295 - 326, DeMoll, E., Escherichia coli and Salmonella, eds. Neidhardt, F. C. et al. ASM Press, Washington DC, USA, 1996: 704 - 708, ISBN 1-55581-084-5).

The synthesis of biotin has been investigated in bacterial organisms, especially in the Gram-negative bacterium Escherichia coli and in the Gram-positive bacterium Bacillus sphaericus (Brown et al. Biotechnol. Genet. Eng. Rev. 9, 1991: 295 - 326). Pimelyl-CoA (PmCoA), which is derived from fatty acid synthesis, has previously been regarded as the first known intermediate in E. coli (DeMoll, E., Escherichia coli and Salmonella, eds. Neidhardt, F. C. et al. ASM Press, Washington DC, USA, 1996: 704 - 708, ISBN 1-55581-084-5 1996). Up to now, the route by which this biotin precursor is synthesized in E. coli has to a large extent been unknown (Lemoine et al., Mol. Microbiol. 19, 1996: 645 - 647). bioC and bioH have been identified as being two genes whose corresponding proteins are responsible for the synthesis of Pm-CoA. The enzymic functions of the gene products, i.e. BioH and BioC, have hitherto been unknown (Lemoine et al., Mol. Microbiol. 19, 1996: 645 - 647, DeMoll, E., Escherichia coli and Salmonella, eds. Neidhardt, F. C. et al. ASM Press, Washington DC, USA, 1996: 704 - 708, ISBN 1-55581-084-5). Pm-CoA is converted into biotin in four further enzymic steps. BioF first of all condenses the Pm-CoA with alanine to form 7-keto-8-aminopelargonic acid (KAPA). The KAPA is then converted into 7,8-diaminopelargonic acid (DAPA)

by BioA. Following an ATP-dependent carboxylation reaction, the next step leads to dethiobiotin (DTB) and is catalyzed by BioD. The DTB is converted into biotin in the last step. This step is catalyzed by BioB. The chemical and enzymic mechanisms involved in the conversion of DTB into biotin are so far only incompletely understood and clarified.

The conversion of DTB into biotin has so far only been characterized in bacterial and plant cell extracts (WO94/8023, EP-B-0 449 724, Sanyal et al. Arch. Biochem. Biophys., Vol. 326, No. 1, 1996: 48 - 56 and Biochemistry 33, 1994: 3625 - 3631, Baldet et al. Europ. J. Biochem. 217, 1, 1993: 479 - 485, Méjean et al. Biochem. Biophys. Res. Commun., Vol. 217, No. 3, 1995: 1231 - 1237, Ohshiro et al., Biosci. Biotechnol. Biochem., 58, 9, 1994: 1738 - 1741).

In vitro studies have demonstrated that low molecular weight factors such as NADPH, cysteine, thiamine, Fe^{2+}, asparagine, serine, fructose 1-6-bisphosphate and S-adenosylmethionine are able to stimulate the synthesis of biotin (Ohshiro et al., Biosci. Biotechnol. Biochem., 58, 9, 1994: 1738 - 1741, Birch et al., J. Biol. Chem. 270, 32, 1995: 19158 - 19165, Ifuku et al., Biosci. Biotechnol. Biochem., 59, 2, 1995: 185 - 189, Sanyal et al. Arch. Biochem. Biophys. 326, 1, 1996: 48 - 56).

In addition to these low molecular weight factors, other proteins have been identified which stimulate the synthesis of biotin from DTB in the presence of BioB. These proteins are flavodoxin and flavodoxin NADPH reductase (Birch et al., J. Biol. Chem. 270, 32, 1995: 19158 - 19165, Ifuk et al., Biosci. Biotechnol. Biochem., 59, 2, 1995: 185 - 189, Sanyal et al., Arch. Biochem. Biophys. 326, 1, 1996: 48 - 56). Other proteins which stimulate biotin synthesis are the genes bioS1 and bioS2, which are described in the German application having the application number 197.31274.8 (Priority 22.7.97).

bioS1 (3) & bioS2 (5) gene

DC 197.31274.8

Differing results have been obtained with regard to the origin of the sulfur in the biotin molecule. Investigations into the synthesis of biotin in whole cell extracts showed that radioactivity was incorporated into biotin in the presence of ^{35}S-labeled cysteine; it was not possible to demonstrate incorporation of sulfur into the biotin molecule when either ^{35}S-labeled methionine or S-adenosylmethionine was used (Ifuku et al., Biosci. Biotechnol. Biochem. 59, 2, 1995: 184 - 189, Birch et al., J. Biol. Chem.270, 32, 1995: 19158 - 19165).

The genes which encode the described proteins, i.e. bioF, bioA, bioD, and bioB, are encoded in E. coli on a bidirectional operon. This operon is located between the λ attachment site and the uvrB gene locus at approx. 17 minutes on the E. coli chromosome (Berlyn et al. 1996: 1715 - 1902). A further two genes, one of which, i.e. bioC, already possesses described functions in the synthesis of Pm-CoA, are additionally encoded on this operon, whereas it has not so far been possible to assign any function to an open reading frame which is located downstream of bioA (WO94/8023, Otsuka et al., J. Biol. Chem. 263, 1988: 19577 - 85). Highly conserved homologues to the E. coli proteins BioF, BioA, BioD and BioB have been found in B. sphaericus, B. subtilis, Syneccocystis sp. (Brown et al. Biotechnol. Genet. Eng. Rev. 9, 1991: 295 - 326, Bower et al., J. Bacteriol. 175, 1996: 4122 - 4130, Kaneko et al., DNA Res. 3, 3, 1996: 109 - 136), archaeobacteria such as Methanococcus janaschi, and yeasts such as Saccharomyces cerevisiae (Zhang et al., Arch. Biochem. Biophys. 309, 1, 1994: 29 - 35) or in plants such as Arabidopsis thaliana (Baldet et al., C. R. Acad. Sci. III, Sci. Vie. 319, 2, 1996: 99 - 106).

In the two Gram-positive microorganisms which have so far been investigated, the synthesis of Pm-CoA appears to proceed in a different manner from that in E. coli. It was not possible to find any homologues of bioH and bioC (Brown et al. Biotechnol. Genet. Eng. Rev. 9, 1991: 295 - 326).

Biotin is an optically active substance which has three centers of chirality. It has hitherto only been prepared economically by way of an expensive, multi-step chemical synthesis.

As an alternative to this chemical synthesis, a large number of attempts have been made to construct a fermentative process for preparing biotin using microorganisms. Cloning the biotin operon onto multi-copy-plasmids has been successfully used to increase biotin synthesis in microorganisms which have been transformed with these genes. A further increase in biotin synthesis was achieved by deregulating biotin gene expression by means of selecting birA mutants (Pai C. H., J. Bacteriol. 112, 1972: 1280 - 1287). Combination of the two approaches, that is expressing the plasmid-encoded biosynthesis genes in a regulation-deficient strain (EP-B-0 236 429), increased productivity still further. In this context, the biotin operon can either remain under the control of its native bidirectional promoter (EP-B-0 236 429) or else its genes can be brought under the control of a promoter which can be regulated externally (WO94/08023).

The approaches which have so far been pursued for producing biotin fermentatively in E. coli have not achieved any economically adequate productivity.

It is an object of the present invention to develop an industrial fermentative process for producing biotin which exhibits as high a biotin synthesis as possible.

We have found that this object is achieved by the process according to the invention for producing biotin, in which process an S-adenosylmethionine synthase (SAM synthase) gene, having the sequence SEQ ID No. 1, and at least one further biotin biosynthesis gene bioS1, bioS2 or bioS3, having the sequences SEQ ID No. 3, SEQ ID No. 5 and SEQ ID No.7, and also their functional variants, analogues or derivatives, are expressed in a prokaryotic or eukaryotic host organism which is able to synthesize biotin, this organism is cultured and the synthesized biotin is used directly after separating off the biomass or after purifying the biotin.

The genes used in the process according to the invention, i.e. the SAM synthase gene having the sequence SEQ ID No. 1 and the biotin biosynthesis genes bioS1, bioS2 and bioS3 having the sequences SEQ ID No. 3, SEQ ID No. 5 and SEQ ID No.7 are kept in the SwissProt-data base under accession numbers P04384 (metK), U29581 (bioS1), P39171 (bioS2) and D90811 (bioS3). A number of homologues to E. coli MetK are described in the data base. These homologues include organisms such as other eubacteria (e.g. H. influenzae, and B. subtilis), and also eukaryotes (e.g. yeasts: S. cerevisiae, Planta: P. deltoides, Arthropoda: D. melanogaster, and Mammalia: R. norvegicus).

The productivity of the biotin biosynthesis can be increased markedly by expressing one or more of the SAM synthase gene, having the sequence SEQ ID No. 1, and its functional variants, analogues or derivatives in combination with at least one of the biotin synthesis genes bioS1, bioS2 or bioS3, having the sequences SEQ ID No. 3, SEQ ID No. 5 and SEQ ID No.7, and also their functional variants, analogues or derivatives, in a prokaryotic or eukaryotic host organism. A combination of the SAM synthase gene and bios1 is preferably used for the expression. At least one further biotin gene selected from the group bioA, bioB, bioF, bioC, bioD, bioH, bioP, bioW, bioX, bioY and bioR is advantageously expressed at the same time in order to increase the biotin synthesis still further. Expression of the genes increases the synthesis of biotin by at least a factor of 2 as compared with the control without these genes, preferably by a

factor which is greater than 3.

The genes used in the process according to the invention, i.e. the SAM synthase gene having the nucleotide sequence SEQ ID No. 1, the bioS1 gene having the nucleotide sequence SEQ ID No. 3, the bioS2 gene having the nucleotide sequence SEQ ID No. 5 and the bioS3 gene having the nucleotide sequence SEQ ID No.7, which sequences encode the amino acid sequences given in SEQ ID NO: 2, SEQ ID No. 4, SEQ ID No. 6 and SEQ ID No.8, respectively, or their allelic variations, can be obtained following isolation and sequencing. Variants are to be understood as being SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and SEQ ID No.7 variants, respectively, which exhibit from 30 to 100% homology at the amino acid level, preferably from 50 to 100% homology, very particularly preferably from 80 to 100% homology. Allelic variants comprise, in particular, functional variants which can be obtained by the deletion, insertion or substitution of nucleotides from the sequences depicted in SEQ ID NO: 1, SEQ ID No. 3, SEQ ID No. 5 and SEQ ID No.7, with, however, the enzymic activity being retained.

In addition, variants are also to be understood as being functional equivalents of the genes, such as O-acetylserine sulfohydrolase A, O-acetylserine sulfohydrolase B, β-cystathionase (see Flint et al., J. Biol. Chem., Vol. 271, 1996: 16053 - 16067) or nifS and its prokaryotic and eukaryotic homologues, for example from Klebsiella, Candida, yeasts or Caenorhabditis, which are able to assume the enzymic activity of bioS1, bioS2 or bioS3 in the synthesis of biotin.

Functional analogues of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID No. 7 are to be understood as being, for example, their prokaryotic or eukaryotic homologues, such as bacterial, fungal, plant, animal or human homologues. In addition, analogues are also to be understood as being truncated sequences, or single-stranded DNA or RNA from coding and non-coding DNA sequences.

Derivatives are to be understood, for example, as being promoter variants. The promoters, which are placed upstream of the given nucleotide sequences, can be altered by means of one or more nucleotide substitutions, or by means of (an) insertion(s) and/or deletion(s) without, however, the functionality or activity of the promoters being impaired. In addition, the activities of the promoters can be increased by means of altering their sequences, or the promoters can be completely replaced by more active

promoters, including those from organisms of a different species.

Derivatives are also to be understood as being variants whose nucleotide sequences have been altered in the region from -1 to -30 upstream of the start codon such that expression of the gene and/or expression of a protein is increased. This is advantageously effected by altering the Shine-Dalgarno sequence.

All Gram-negative or Gram-positive bacteria which synthesize biotin are, in principal, suitable for use as prokaryotic host organisms in the process according to the invention. Gram-negative bacteria which may be mentioned by way of example are Enterobacteriaceae such as the genera Escherichia, Aerobacter, Enterobacter, Citrobacter, Shigella, Klebsiella, Serratia, Erwinia or Salmonella, Pseudomonadaceae such as the genera Pseudomonas, Xanthomonas, Burkholderia, Gluconobacter, Nitrosomonas, Nitrobacter, Methanomonas, Comamonas, Cellulomonas or Acetobacter, Azotobacteraceae such as the genera Azotobacter, Azomonas, Beijerinckia or Derxia, Neisseriaceae such as the genera Moraxella, Acinetobacter, Kingella, Neisseria or Branhamella, the Rhizobiaceae such as the genera Rhizobium or Agrobacterium, or the Gram-negative genera Zymomonas, Chromobacterium or Flavobacterium. Gram-positive bacteria which may be mentioned by way of example are the endospore-forming Gram-positive aerobic or anaerobic bacteria such as the genera Bacillus, Sporolactobacillus or Clostridium, the coryneform bacteria such as the genera Arthrobacter, Cellulomonas, Curtobacterium, Corynebacterium, Brevibacterium, Microbacterium or Kurthia, the Actinomycetales such as the genera Mycobacterium, Rhodococcus, Streptomyces or Nocardia, the Lactobacillaceae such as the genera Lactobacillus or Lactococcus, or the Gram-positive cocci such as the genera Micrococcus or Staphylococcus.

Preference is given to using bacteria of the genera Escherichia, Citrobacter, Serratia, Klebsiella, Salmonella, Pseudomonas, Comamonas, Acinetobacter, Azotobacter, Chromobacterium, Bacillus, Clostridium, Arthrobacter, Corynebacterium, Brevibacterium, Lactococcus, Lactobacillus, Streptomyces, Rhizobium, Agrobacterium or Staphylococcus in the process according to the invention. Particular preference is given to genera and species such as Escherichia coli, Citrobacter freundii, Serratia marcescens, Salmonella typhimurium, Pseudomonas mendocina, Pseudomonas aeruginosa, Pseudomonas mutabilis, Pseudomonas chlororaphis, Pseudomonas fluorescens, Comamonas acidovorans, Comamonas testosteroni, Acinetobacter calcoaceticus, Azotobacter vinelandii, Chromobacterium violaceum, Bacillus subtilis, Bacillus sphaericus, Bacillus stearothermophilus, Bacillus

pumilus, Bacillus licheniformis, Bacillus amyloliquefaciens, Bacillus megaterium, Bacillus cereus, Bacillus thuringiensis, Arthrobacter citreus, Arthrobacter paraffineus, Corynebacterium glutamicum, Corynebacterium primorioxydans, Corynebacterium sp., Brevibacterium ketoglutamicum, Brevibacterium linens, Brevibacterium sp., Streptomyces lividans, Rhizobium leguminosarum or Agrobacterium tumefaciens. Advantageously, use is made of bacteria which already exhibit an elevated natural production of biotin.

The taxonomic position of the listed genera has been subject to considerable change in recent years and is still in a state of flux as false genera and species names are corrected. Because of these taxonomic regroupings, which have been frequently required in the past, of the said genera within bacterial systematics, families, genera and species other than those mentioned above are also suitable for the process according to the invention.

All biotin-synthesizing organisms, such as fungi, yeasts, plants or plant cells, are, in principal, suitable for use as eukaryotic host organisms in the process according to the invention. Yeasts which may preferably be mentioned are the genera Rhodotorula, Yarrowia, Sporobolomyces, Saccharomyces or Schizosaccharomyces. Particular preference is given to the genera and species Rhodotorula rubra, Rhodotorula glutinis, Rhodotorula graminis, Yarrowia lipolytica, Sporobolomyces salmonicolor, Sporobolomyces shibatanus or Saccharomyces cerevisiae.

In principal, all plants can be used as the host organism, with preference being given to plants which play a role in animal nutrition or human nutrition, such as corn, wheat, barley, rye, potatoes, peas, beans, sunflowers, palms, millet, sesame, copra or rape. Plants such as Arabidopsis thaliana or Lavendula vera are also suitable. Particular preference is given to plant cell cultures, plant protoplasts or callus cultures.

Microorganisms such as bacteria, fungi, yeasts or plant cells which are able to secrete biotin into the growth medium, and which, where appropriate, already additionally exhibit an increased natural synthesis of biotin, are advantageously used in the process according to the invention. Advantageously, these organisms can also be defective with regard to the regulation of their biotin biosynthesis; i.e. this synthesis is either not regulated or only regulated to a very reduced extent. This regulatory defect results in these organisms already possessing a substantially increased biotin productivity. Such a regulatory defect is known, for example, from Escherichia coli in the form

of birA-defect mutants and should preferably be present in the cells as a defect which can be induced by external influences, for example as a defect which is temperature-inducible. In principal, organisms which do not exhibit any natural biotin production can also be used, once they have been transformed with the biotin genes.

In order to increase biotin productivity as a whole still further, the organisms in the process according to the invention should advantageously also harbor at least one further biotin gene selected from the group bioA, bioB, bioF, bioC, bioD, bioH, bioP, bioW, bioX, bioY or bioR. Advantageously, those genes which stimulate biotin synthesis can also be present in the cell in combination with the sequences SEQ ID No. 1 , SEQ ID No. 3, SEQ ID No.5 or SEQ ID No.7 and their combinations. Examples of genes which stimulate biotin synthesis are the flavoredoxin gene and the flavoredoxin reductase gene. This additional gene, or these additional genes, can be present in the cell in one or more copies, like the genes having the sequences SEQ ID No. 1 , SEQ ID No.3, SEQ ID No.5 or SEQ ID No.7 or their combinations. They can be located on the same vector as the sequences SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No.7, or on separate vectors, or else integrated chromosomally. The sequences SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No.7 can also be together on a vector or on separate vectors or be inserted into the genome.

The gene construct according to the invention is to be understood as being the gene sequences of the SAM synthase gene SEQ ID No. 1 and of the biotin synthesis genes SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No.7, and also their functional variants, analogues or derivatives, which were linked functionally to one or more regulatory signals for the purpose of increasing expression of the genes. In addition to these new regulatory sequences, the natural regulation of these sequences can still be present upstream of the actual structural genes and, where appropriate, can have been genetically altered such that the natural regulation has been switched off and expression of genes has been increased. However, the gene construct can also be assembled in a simpler manner, i.e. no additional regulatory signals are inserted upstream of the sequences SEQ ID No. 1, SEQ ID No. 3, SEQ ID No.5 and/or SEQ ID No.7 and the natural promoter, with its regulation, is not removed. Instead, the natural regulatory sequence is mutated such that regulation by biotin no longer takes place and gene expression is increased. The sequences SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No.7 can be under the regulation of one promoter or under the regulation of separate promoters. Additional, advantageous regulatory elements can also be inserted at the 3' end of the DNA sequences. The

genes having the sequences SEQ ID No. 1, SEQ ID No. 3, SEQ ID No.5 or SEQ ID No. 7 can be present in the gene construct in one or more copies.

Advantageous regulatory sequences for the process according to the invention are present, for example, in promoters such as the cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, $lacI^q$-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, λ-P_R- or λ-P_L-promoters, which are advantageously used in Gram-negative bacteria. Further advantageous regulatory sequences are present, for example, in the Gram-positive promoters amy and SPO2, in the yeast promoters ADC1, MFα , AC, P-60, CYC1, GAPDH or in the plant promoters CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, or nos, or in the ubiquitin promoter or the phaseolin promoter.

In principal, all natural promoters, together with their regulatory sequences, can be used, like the abovementioned promoters, for the process according to the invention. In addition, synthetic promoters can also advantageously be used.

Other biotin genes selected from the group bioA, bioB, bioF, bioC, bioD, bioH, bioP, bioW, bioX, bioY or bioR, which genes can have their own promoter or else can be under the regulation of the promoter of one of the sequences, or under the regulation of the promoter of all the sequences, SEQ ID No. 1, SEQ ID No. 3, SEQ ID No.5 or SEQ ID No.7, can be present in the gene construct in one or more copies.

For expression in the abovementioned host organism, the gene construct is advantageously inserted into a host-specific vector which makes it possible to achieve optimum expression of the genes in the host. Vectors are well known to the skilled person and can be identified, for example, from the book Cloning Vectors (Eds. Pouwels P. H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018). In addition to plasmids, the vectors are also to be understood as being all other vectors known to the skilled person, such as phages, viruses, transposons, IS elements, phasmids, cosmids or linear or circular DNA. These vectors can be replicated autonomously in the host organism or replicated chromosomally.

Expression systems are to be understood as being the combination of the host organisms which are mentioned above by way of example and the vectors which are appropriate for the organisms, such as plasmids, viruses or phages, for example plasmids containing the RNA polymerase/promoter system, phages λ, or Mu or other temperate phages or transposons and/or further advantageous regulatory

sequences.

The term expression systems is preferably to be understood as being the combination of Escherichia coli and its plasmids and phages and the affiliated promoters, and also Bacillus and its plasmids and promoters.

Further 3' and/or 5'-terminal regulatory sequences are also suitable for advantageously expressing SEQ ID No.1, SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No. 7 in accordance with the invention.

These regulatory sequences are intended to make it possible to achieve specific expression of the biotin genes and expression of the protein. Depending on the host organism, this can, for example, mean that the gene is only expressed or overexpressed after induction or that it is expressed and/or overexpressed immediately.

In this context, the regulatory sequences or factors can preferably influence biotin gene expression positively and thereby increase it. For example, the regulatory elements can advantageously be reinforced at the transcriptional level by means of using strong transcription signals such as promoters and/or enhancers. In addition, however, it is also possible to reinforce translation by, for example, improving the stability of the mRNA.

Enhancers are to be understood as being, for example, DNA sequences which bring about increased biotin gene expression by means of improving the interaction between the RNA polymerase and the DNA.

An increase in the proteins (see SEQ ID No.2, SEQ ID No.4, SEQ ID No.6 and SEQ ID No.8) which are derived from the sequences SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and SEQ ID No.7, and in their enzyme activity, as compared with the starting enzymes, can be achieved, for example, by altering the corresponding gene sequences, or the sequences of their homologues, by means of classical mutagenesis, such as UV irradiation, or by treating with chemical mutagens and/or by means of specific mutagenesis such as site-directed mutagenesis, deletion(s), insertion(s) and/or substitution(s). An increased enzyme activity, apart from the described gene amplification, can also be achieved by eliminating factors which repress enzyme biosynthesis and/or by synthesizing active enzymes instead of inactive enzymes.

The process according to the invention advantageously increases the conversion of DTB into biotin, and consequently overall biotin productivity, by means of using the biotin genes having the sequences SEQ ID No. 1, SEQ ID No.3, SEQ ID No.5 and SEQ ID No.7, and the combination of the genes having the sequences SEQ ID No.1 and SEQ ID No.5 or SEQ ID No.1 and SEQ ID No.7, preferably the combination of the genes having the sequences SEQ ID No.1 and SEQ ID No.3, which genes are introduced into the organisms by way of their vectors and/or by means of chromosomal cloning.

In the process according to the invention, the microorganisms harboring SEQ ID No.1, SEQ ID No.3, SEQ ID No.5 and/or SEQ ID No.7 are propagated in a medium which enables these organisms to grow. This medium can be a synthetic medium or a natural medium. Use is made of media which are known to the skilled person and which are appropriate for the organism. In order to permit growth of the microorganisms, the media employed contain a carbon source, a nitrogen source, inorganic salts and, where appropriate, small quantities of vitamins and trace elements.

Examples of advantageous carbon sources are sugars, such as monosaccharides, disaccharides or polysaccharides, such as glucose, fructose, mannose, xylose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose, complex sugar sources such as molasses, sugar phosphates, such as fructose-1,6-bisphosphate, sugar alcohols, such as mannitol, polyols, such as glycerol, alcohols, such as methanol or ethanol, carboxylic acids, such as citric acid, lactic acid of acetic acid, fats, such as soy-bean oil or rape-seed oil, or amino acids, such as glutamic acid or aspartic acid, or amino sugars, which can simultaneously be used as a nitrogen source.

Advantageous nitrogen sources are organic or inorganic nitrogen compounds or materials which contain these compounds. Examples are ammonium salts, such as NH_4Cl or $(NH_4)_2SO_4$, nitrates or urea, or complex nitrogen sources such as corn steep liquor, brewer's yeast autolysate, soy-bean flour, wheat gluten, yeast extract, meat extract, casein hydrolysate or yeast or potato protein, which can frequently also be used simultaneously as a nitrogen source.

Examples of inorganic salts are the salts of calcium, magnesium, sodium, manganese, potassium, zinc, copper and iron. Anions of these salts which are to be mentioned in particular are the chloride, sulfate and phosphate ions. An important factor for

increasing productivity in the process according to the invention is the addition of Fe^{2+} or Fe^{3+} salts and/or potassium salts to the production medium.

Where appropriate, further growth factors, such as vitamins or growth promoters, such as riboflavin, thiamine, folic acid, nicotinic acid, pantothenate or pyridoxine, amino acids, such as alanine, cysteine, asparagine, aspartic acid, glutamine, serine, methionine or lysine, carboxylic acids, such as citric acid, formic acid, pimelic acid or lactic acid, or substances such as dithiothreitol, are added to the nutrient medium.

Antibiotics can, where appropriate, be added to the medium in order to stabilize the biotin gene-containing vectors in the cells.

The ratios in which the said nutrients are mixed depends on the nature of the fermentation and is laid down in each individual case. The medium components can all be initially introduced at the beginning of fermentation, after they have been, if necessary, sterilized separately or sterilized together, or else be added subsequently, as required, during fermentation.

The culture conditions are so arranged that the organisms grow optimally and that the best possible yields are achieved. Preferred culture temperatures are from 15 °C to 40 °C. Temperatures of between 25 °C and 37 °C are particularly advantageous. The pH is preferably kept in a range of from 3 to 9. pH values of between 5 and 8 are particularly advantageous. In general, a period of incubation of from 8 to 240 hours, preferably of from 8 to 120 hours, is sufficient. Within this time, the maximum quantity of biotin accumulates in the medium and/or is available after the cells have been disrupted.

The process according to the invention for producing biotin can be carried out continuously or batch-wise or fed-batch-wise. If whole plants are regenerated from the plant cells which have been transformed with the biotin genes, they can, according to the process according to the invention, be grown and propagated perfectly normally.

Examples:

1. Cloning of the S-adenosylmethionine synthase gene (SEQ ID No.1):

Starting from genomic E. coli DNA, the gene which encodes SAM synthase (metK) was amplified from the E. coli chromosome by means of a polymerase chain reaction using two specific oligonucleotides. The DNA which had been amplified in this way was purified, digested with the restriction enzyme Acc65I and inserted into a vector which had been cut with the same enzyme and which enables the gene to be overexpressed in E. coli strains. One of the two oligonucleotides was used to provide the gene construct with optimized translation signals.

a.) Generation of oligonucleotides for amplifying the metK gene from the E. coli chromosome:

metK was to be amplified as an expression cassette which was composed of a ribosome binding site, the start codon of the coding sequence and the stop codon between two restriction enzyme recognition sites. The Acc65I recognition sequence was chosen for both the restriction sites. The metK gene was amplified and cloned using the nucleotides PmetK1 (5'-GCGGTACCAGGTGATATTAAATATGGCAAAAC-3') and PmetK2 (5'-CGGGTACCGATTACTTCAGACCGGCAGC-3').

b.) Implementation of the PCR:

Conditions:

0.5 µg chromosomal DNA from E. coli W3110 was used as a template. The oligonucleotides PmetK1 and PmetK2 were employed at a concentration of in each case 15 pMol. The concentration of the dNTPs was 200 µM. 2.5 U of Pwo DNA polymerase (Boehringer Mannheim) in the manufacturer's reaction buffer were employed as the polymerase. The PCR reaction volume was 100 µl.

Amplifications:

The DNA was denatured at 94 °C for 2 min. The oligonucleotides were then annealed at 55 °C for 30 seconds. The elongation took place at 72 °C for 75 seconds. The PCR reaction was carried out over 30 cycles.

The resulting DNA product, which had a size of approximately 1145 bp, was purified and digested with Acc65I in a suitable buffer.

c.) Cloning of metK in an expression vector

2 µg of the vector pHS1 (construction was described in DE 197.31274.8, priority 22.7.97, Example 1, pages 14 to 17) were digested with Acc65I and dephosphorylated using shrimp alkaline phosphatase (SAP) (Boehringer Mannheim). After the SAP had been denatured, vector and fragment were ligated, in a molar ratio of 1:3, using the Rapid DNA Ligation kit in accordance with the manufacturer's instructions. The ligation mixture was transformed into strain E. coli XL-1-blue. Positive clones were identified by plasmid preparation and restriction analysis. The correct orientation of the metK fragment in pHS1 was determined by restriction digestion and sequencing. The resulting construct was designated pHS1 metK (Figure 1). The sequence of pHS1 metK is given in SEQ ID No.9. SEQ ID No.10 shows the amino acid sequence which is deduced from the metK-encoding region.

2. Construction of plasmids pHBbio14 and pHS1 bioS1

The construction of plasmids pHBbio14 and pHS1 bioS1 has already been described (DE 197.31274.8, Priority 22.7.97, Examples 1, 2 and 5).

3. Construction of pHS1 metK bioS1

The plasmids pHS1 bioS1 [SEQ ID No.11, (DE 197.31274.8, Priority 22.7.97), SEQ ID No.12 shows the amino acid sequence which is deduced from the bioS1-encoding region] and pHS1 metK (SEQ ID No.9) were purified using a plasmid preparation method (Boehringer). The fragment carrying the metK gene was isolated from pHS1 metK by digesting with Acc65I. pHS1 bioS1 was digested with Acc65I and dephosphorylated with shrimp alkaline phosphatase (SAP) (Boehringer Mannheim). After the SAP had been denatured in accordance with the manufacturer's instructions, the vector and the metK fragment were ligated, in a molar ratio of 1:3, using the Rapid DNA Ligation Kit in accordance with the manufacturer's instructions. The ligation mixture was transformed into strain E. coli XL-1-blue. Positive clones were identified by plasmid preparation and restriction analysis. The correct orientation of the metK fragment in pHS1 bioS1 was determined by means of restriction digestion and sequencing. The resulting construct was designated pHS1 metK bioS1 (Figure 2). The sequence of pHS1 metK bioS1 is given in SEQ ID No.13. SEQ ID No.14 shows the amino acid sequence which was deduced from the metK-encoding region; SEQ ID

No.15 shows the amino acid sequence which was deduced from the bioS1-encoding region.

4. Increasing biotin productivity by overexpressing metK, bioS1 and metK in combination with bioS1.

Spontaneously rifampicin-resistant colonies were isolated from strain BM4086 (Ketner and Campbell J. Molec. Biology 1975 96:13) by plating on rifampicin plates. A P1 lysate was generated from one of these resistant strains. The strain W3110 was transduced with this P1 lysate and clones were subsequently selected using rifampicin. The resulting strain was transformed with plasmid pHBbio14 using the $CaCl_2$ method (Maniatis et al. Molecular Cloning Cols Spring Harbour Laboratory Press 1989) and grown on LB containing 100 µg of ampicillin/ml. The isolated, transformed strain (LU5560) was in each case transformed with plasmid pHS1, pHS1 metK, pHS1 bioS1 or pHS1 metK bioS1 using the $CaCl_2$ method and then selected on LB agar containing 100 µg of ampicillin/ml and 25 µg of kanamycin/ml.

One colony from each of the transformants was in each case inoculated into a DYT culture containing the appropriate antibiotics and incubated for 12 h. The overnight culture (= ONC) was used to inoculate a 10 ml culture in TB medium (Sambrook, J. Fritsch, E F. Maniatis, T. 2nd ed. Cold Spring Harbor Laboratory Press., 1989 ISBN 0-87969-373-8), which contained 30 g of glycerol/l and the appropriate antibiotics. In the cases where plasmids pHS1, pHS1 metK, pHS1bioS1 and pHS1 metK bioS1 were present, 1mM IPTG and 0.5% arabinose were added simultaneously in order to induce expression of the metK and bioS1 genes or, respectively, the combination of the two genes. After 24 h, the cells were separated off from the culture supernatant by centrifugation and the biotin concentration in the supernatant was determined by means of a competitive ELISA employing streptavidin. The results of this determination are shown in Table I.

Table I: Determination of the biotin concentration

Strain	Plasmid I	Plasmid II	Biotin mg/l
LU5580	pHBbio14	Control, without plasmid	11
LU5580	pHBbio14	pHS1	25
LU5580	pHBbio14	pHS1 bioS1	45
LU5580	pHBbio14	pHS1 metK	37
LU5580	pHBbio14	pHS1 metK bioS1	52